A regression method with subnetwork neurons for vigilance estimation using EOG and EEG

Wei Wu, *Student Member, IEEE,* Q. M. Jonathan Wu, *Senior Member, IEEE,* Wei Sun, Yimin Yang, *Member, IEEE*, Xiaofang Yuan, *Member, IEEE,* Wei-Long Zheng, *Student Member, IEEE* and Bao-Liang Lu, *Senior Member, IEEE*

***Abstract*—In recent years, it has been observed that there is an increasing rate of road accidents due to the low vigilance of drivers. Thus, the estimation of drivers' vigilance state plays a significant role in Public Transportation Safety (PTS). We have adopted a feature fusion strategy that combines the electroencephalogram (EEG) signals collected from various sites of the human brain, including forehead, temporal, and posterior and forehead electrooculogram (forehead-EOG) signals, to address this factor. The level of vigilance is predicted through a new learning model known as double-layered neural network with subnetwork nodes (DNNSN), which comprises several subnetwork nodes, and each node in turn is composed of many hidden nodes that have various capabilities of feature selection (dimension reduced), feature learning, etc. The proposed single modality that uses only forehead-EOG signal exhibits a mean root-mean-square error (RMSE) of 0.12 and a mean Pearson product-moment correlation coefficient (COR) of 0.78. On one hand, an EEG signal achieved a mean RMSE of 0.13 and a mean COR of 0.72. Whereas, on the other, the proposed multimodality achieved values of 0.09 and 0.85 for the mean RMSE and the mean COR, respectively. Experimental results show that the proposed DNNSN with multimodality fusion outperforms the model with single modality for vigilance estimation due to the complementary information between forehead-EOG and EEG. After a favorable learning rate was applied to the input layer, the mean RMSE/COR improved to 0.11/0.79, 0.12/0.74, and 0.08/0.86, respectively. Hence, this quantitative analysis proves that the proposed method provides better feasibility and efficiency learning capability and surmounts other state-of-the-art techniques.**

***Index Terms*—EEG, forehead-EOG, feedforward neural network, learning rate, vigilance estimation.**

This work was supported in part by the Canada Research Chair Program and the NSERC Discovery Grant; by the Independent Research Project of State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body 71765003, and the Hunan Key Laboratory of Intelligent Robot Technology in Electronic Manufacturing Open Foundation of China under Grant 2017TP1011; by the National Key Research and Development Program of China under Grant 2017YFB1002501, and the Fundamental Research Funds for the Central Universities; by the National Natural Science Foundation of China under Grant 61673266, 61573133 and 61401046.
(Corresponding authors: Wei Sun; Q. M. Jonathan Wu; Yimin Yang.)

W. Wu, W. Sun, and X. F. Yuan are with the College of Electrical and Information Engineering, Hunan University, the State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, and the Hunan Key Laboratory of Intelligent Robot Technology in Electronic Manufacturing, Hunan University, Changsha 410082, China.

Q. M. Jonathan Wu is with the Department of Electrical and Computer Engineering, University of Windsor, Windsor N9B 3P4, Canada.

Y. M. Yang is with Computer Science Department, Lakehead University, Thunder Bay P7B 5E1, Canada.

W. L. Zheng and B. L. Lu are with the Department of Computer Science and Engineering, Shanghai Jiao Tong University and the Key Laboratory of Shanghai Education Commission for Intelligent Interaction and Cognitive Engineering, Shanghai Jiao Tong University, Shanghai 200240, China.

I. Introduction

Vigilance is a vital physiological signal and usually defined as the ability of organisms to maintain their long-term attention for stimuli [1]. When people engaging in public transportation (e.g., pilots, drivers, etc.) reduce or lose their vigilance, fundamentally serious accidents occur, sometimes resulting in fatal events. New South Wales (NSW) [2] investigated that of all the fatal traffic accidents from 2009 to 2013, at least 20% were caused by fatigue driving. In 2000, the National Aeronautics and Space Administration (NASA) [3] reported that a famous questionnaire was set for almost 11,000 pilots from 2,000 companies, of which over half-a-dozen surveys indicated that pilots were exhibited with fatigue during the approach/landing flight phase. Fatal traffic accidents due to reduced vigilance have shown to become more and more common around the world in recent years [4], [5].

To estimate the level of vigilance, typically, methods can be divided into four categories [6]: physiological methods [7]–[27], behavioral methods [28]–[34], subjective methods [35]–[38], and vehicle-based methods [39]–[41].

A. Physiological Methods

Most physiological measures are based on the detection of drowsiness using mainly the following four procedures:

1) Electroencephalogram (EEG): Loomis *et al.* [7] indicated that electrical waves correlated with physiological or psychological states are recorded from electric potentials of the human brain to analyze the difference between the states of wakefulness and sleep where the electroencephalogram (EEG) is used first in order to estimate vigilance. An experimental analysis of 22 subjects showed that the slower the wave, the deeper the sleep. The quantitative estimation of the level of alertness such as testing blood alcohol content (BAC) was proposed in previous works [8]–[10]. Matousek *et al.* [8] introduced an EEG-based multi-regression model to automatically determine the mean Pearson product-moment correlation coefficient (COR) through one set of data that can also be used for the spectral values of another set of data. According to an activity index, which was computed by beta to delta ratio and beta and delta components of EEG features, Merica *et al.* [9] showed that values of gradient and magnitude rapidly changed at the beginning or just after stage 1 (drowsy) and got an accuracy of 75%. Makeig *et al.* [10] used a two-fold coherence experiment for analyzsing fluctuations to evaluate the linear relationship between the extracted EEG signals by moving-average measure and levels of vigilance, and then, verifying the linear correlation by strong converging experimental results. Khushaba *et al.* [11] used the fuzzy mutual-information-based wavelet-packet transform (FMIWPT) algorithm to extract features from the datasets collected from 31 drivers in an experimental environment, and this algorithm showed an accuracy of at least 95%. Recently, various advanced learning techniques have been successfully applied in areas of speech recognition [42], image recognition [43], and cognitive computing, etc. For example, Yang *et al.* [13] used the subnetwork nodes of a hierarchical network to recognize three emotions of negative, neutral, and positive. Every subnetwork node as a hidden layer independently extracted the subspace features, and the classification of accuracy improved to 91%. Chai *et al.* [14] obtained the experimental samples that were evaluated from 43 subjects between alert and tired states. Then, they used the autoregressive (AR) model and the sparse-deep belief networks (s-DBNs) [44] as the features extraction and the classification algorithms, respectively, and improved the sensitivity to 93%. It was noted that EEG directlyrecorded neurophysiological signals that were correlated with alertness; thus, it can be considered a reliable method for vigilance estimation.

2) Electrooculogram (EOG): Electrooculogram (EOG) signals contain essential information from various eye movements. Wierwille *et al.* proposed a famous algorithm known as the percentage of eyelid closure (PERCLOS) using a high-resolution camera to test eye closure over 80% to judge drowsiness. Many commercialized fatigue driving devices, with the PERCLOS algorithm [15], gradually appeared and were ultimately recognized by the National Highway Traffic Safety Administration (NHTSA) [16]. Hu *et al.* [17] utilized support vector machine (SVM) strategy to obtain a mean accuracy of 90% by using EOG and EMG signals recorded from six electrodes and one electrode, respectively. Compared to the traditional EOG signal, which was collected from the traditional electrode placement, Zhang *et al.* [18] obtained a new EOG signal that was collected from the forehead electrode placement and used eye-tracking glasses [19] to calculate the PERCLOS index. Then, they used the SVM algorithm to estimate vigilance and obtain a high COR of 0.86. Ma *et al.* [20] reported that an EOG signal has two critical characteristics: an easy setup and a high signal-to-noise ratio. Huo and Zheng *et al.* [21], [22] found the complementarity of forehead-EOG and EEG signals, and using the feature fusion for vigilance estimation, improved the accuracy rate. In short, EOG-based methods gradually played a vital role for vigilance estimation.

3) Electrocardiogram (ECG) and Electromyogram (EMG): In recent years, many methods have used Electrocardiogram (ECG) [11], [23]–[25] and Electromyogram (EMG) [26], [27] signals to quantitatively analyze levels of vigilance. For example, Patel *et al.* [24] proposed a strategy that utilized the bandpass filter (BPF) for pre-processing raw ECG signal and extracted features by fast Fourier transform (FFT). They then used a neural network (NN) algorithm, obtaining an accuracy of 90%. Compared to the experimental results between wakefulness and drowsiness, Meng *et al.* [25] however found that subjects' blood pressure and respiratory rate, instead of heart rate, had changed significantly. Boonleng *et al.* [27] developed a mobile system consisting of five wearable sensors that were placed in the best positions to detect vigilance. Then, SVM-based estimation vigilance algorithm showed a detection accuracy rate of 92%. In general, ECG and EMG-based methods for vigilance estimation obtain good results, but there are still many technical challenges.

Despite of this, the physiological-based method can be considered as an effective and objective measure of levels of vigilance.

B. Behavioral Methods

Behavioral-based vigilance estimation methods use features that contain mouth states [28], [29], eye states [30]–[32], facial expressions [33], and body posture [34] collected by a video device (e.g., camera, Infrared illuminator) to compute the detection accuracy rate. Alioua *et al.* [28] proposed the circular Hough transform (CHT)-based approach using mouth state feature detected by a circular edge, which showed the mean correct classification rate (MCCR) and kappa statistic (K) as 0.98 and 0.97, respectively. In addition, Flores *et al.* [31] proposed a support vector machine (SVM)-based model with eye state features, which were extracted using a condensation algorithm, with an accuracy of 93%. Orazio *et al.* [32] proposed a neural classifier using eye state feature extraction by discrete wavelet transform (DWT) and then a Hough transform (HT)-based method to achieve an accuracy of 95%. Murphy *et al.* [34] designed a new system with a particle filter, which combined 3-d face models of a support vector regressors (SVRs) [45]-based approach, by utilizing local-oriented gradients (LOG) to recognize the static head posture. They then, under laboratory conditions of daily driving and night driving, obtained a mean absolute error of pitch (MAEP) and a mean absolute error of yaw (MAEY) of a static pose as $4.92^o/7.81^o$ and $1.64^o/2.08^o$, respectively. One of the limitations of these methods is their ignorance of the uniqueness of different characteristics and habits of each driver.

C. Subjective Methods

Subjective-based methods reported in literature mainly include seven-point Stanford sleepiness scale (SSS) [35], visual analogue scales (VAS) [36], Epworth sleepiness scale (ESS) [37], nine-point Karolinska sleepiness scale (KSS) [38], etc. The ESS is widely used for evaluating levels of sleepiness, and each score is obtained from 8 different daily questions to assess the probability of falling asleep. The range of values, '0-9', '10-15', and '16-24' represent normal, moderate sleep apnea, and severe sleep apnea, respectively. The primary limitation of these methods is that they are self-reported measures that are based on personal biases.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2018.2889223, IEEE Transactions on Cognitive and Developmental Systems

D. Vehicle-Based Methods

It has already been established that most of the previous vehicle-based approaches have been taken from literature [39]–[41]. Researchers mainly focus on the following features: average measured distance from lane center, standard analysis deviation of speed and lane position, measurement of the mean vehicle speed from speed limit, and monitoring of the lane position and the angle of the steering wheel. Basically, there are 10-22 subjects employed for driving tasks such as more than 30-mins of drive alone on a simulated road by placing sensors and lasting at least 16 hours or more without sleep or mainly depending on drugs such alcohol or caffeine to obtain drowsiness. For example, He *et al.* [40] proposed Bayesian network (BN)-based vigilance estimation approach using the experimental data built by EEG from 10 subjects to compute the lane deviation on simulated environmental conditions, which showed that the COR is around 0.05. Compared with a monotonous simulated environment, the real road condition is more complicated, including the increase in lateral distance with speed, leading to higher a risk; this did not appear in most experiments.

From the methods explained above, two machine learning methods can be mainly used for an automated prediction of the vigilance: the classification method and the regression method. The goal of the classification algorithms is to predict the subject, whether in the state of fatigue or alert, while the outcome of regression algorithms is to predict continuous values for vigilance estimation.

According to the complementarity of forehead-EOG and EEG, we proposed a new learning model known as the double-layered neural network with subnetwork nodes (DNNSN) using feature fusion to improve the accuracy of prediction of vigilance estimation. In particular, the paper contributed as follows:

1) We used the forehead-EOG and EEG signals that were collected from the SEED-VIG dataset[1]. To evaluate vigilance, we proposed the DNNSN model comprising of many hidden nodes that have various capabilities of feature selection (dimension reduced), feature learning [46], [47], etc., and then, the promising experimental results consistent with previous studies [48]. When we used the learning rate in the entrance layer, the accuracy of prediction significantly improved.

2) The DNNSN model can be applied to all physiological signals, whether single modality or multimodality. For single modality, the mean RMSE/COR of the proposed method using forehead-EOG and EEG features are 0.12/0.78 and 0.13/0.72, respectively. For multimodality, the mean RMSE/COR of the proposed method using the feature fusion is 0.09/0.85. Furthermore, after we utilized the learning rate in the entrance layer, through forehead-EOG, EEG, and the feature fusion, the results of the proposed method improved to 0.11/0.79, 0.12/0.74, and 0.08/0.86, respectively, which is outperformed other state-of-the-art methods.

[1]http://bcmi.sjtu.edu.cn/~seed/

3) We found the proposed method using EOG-ICA-MINUS, which has a good result consistent with previous studies [12], [22]. This is because the independent component analysis (ICA) approach can easily detect blink components, such as impulses from vertical EOG (VEO) features, and the MINUS approach can easily identify saccade components from Horizontal EOG (HEO) features. We also observed that this method has a better performance in comparison to the EOG-ICA-MINUS. Strictly speaking, the results are very close. We demonstrated that the proposed method has better performance in detecting saccade, blink, and fixation components. Here, VEO_f-ICA represents forehead-VEO features extracted by ICA; HEO_f-MINUS represents forehead-HEO features extracted by MINUS and the subscribe 'f' represents forehead.

4) The standard deviation (STD) of a data set reflects its degree of dispersion, and the smaller this value, the lesser is the deviation from the average, and vice versa. The proposed method can provide the lowest of the STD of RMSE and COR, which explains that our results are better than those of other comparison methods.

In general, the proposed method can be considered as a robust regression model to estimate levels of vigilance for the multimodality by using the fused features.

II. Methodology

As mentioned earlier [20], the forehead-EOG signals have characteristics of easy setup and high signal-to-noise ratio and contain interference noise with blink, saccade, and fixation components. EEG signals should also be considered as a trustworthy method for vigilance estimation, as they directly record neurophysiological signals that are correlated with alertness. The complementary of the two signals is a major focal point for vigilance estimation and, based on this characteristic, we proposed a multilayer network with the subnetwork nodes using the fused signals. Simultaneously, we used the learning rate for the entrance layer to extract subspace features from the input data. With reduced dimensions of the input data of the entrance layer, the output data of the exit layer are fused gather for the final regression analysis.

A. Network Model

The structure of the proposed method can be seen in Fig. 1. Here, both yellow dots of the entrance layer and red dots of the exit layer represent hidden nodes; blue dots represent the subnetwork nodes. The input data x represents forehead-EOG or EEG signals. The output data y represents awake, tired, and drowsy states in the range of '0' to '1'. Based on the PERCOLS, we get two thresholds of '0.35' and '0.70', and ranges '0-0.35', '0.36-0.70', and '0.71-1' represent awake, tired, and drowsy states, respectively. The symbols used for the proposed method are defined in Table I. The process of subspace features extraction and combination are described as follows:

Step (1): For the entrance layer, given $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{M}, \mathbf{x}_i \in \mathbb{R}^m$ arbitrary distinct training samples from a continuous system, the weight $(\hat{\omega}_p^k)$, and the bias $(\hat{b}_p^k)$ are obtained by the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2018.2889223, IEEE Transactions on Cognitive and Developmental Systems

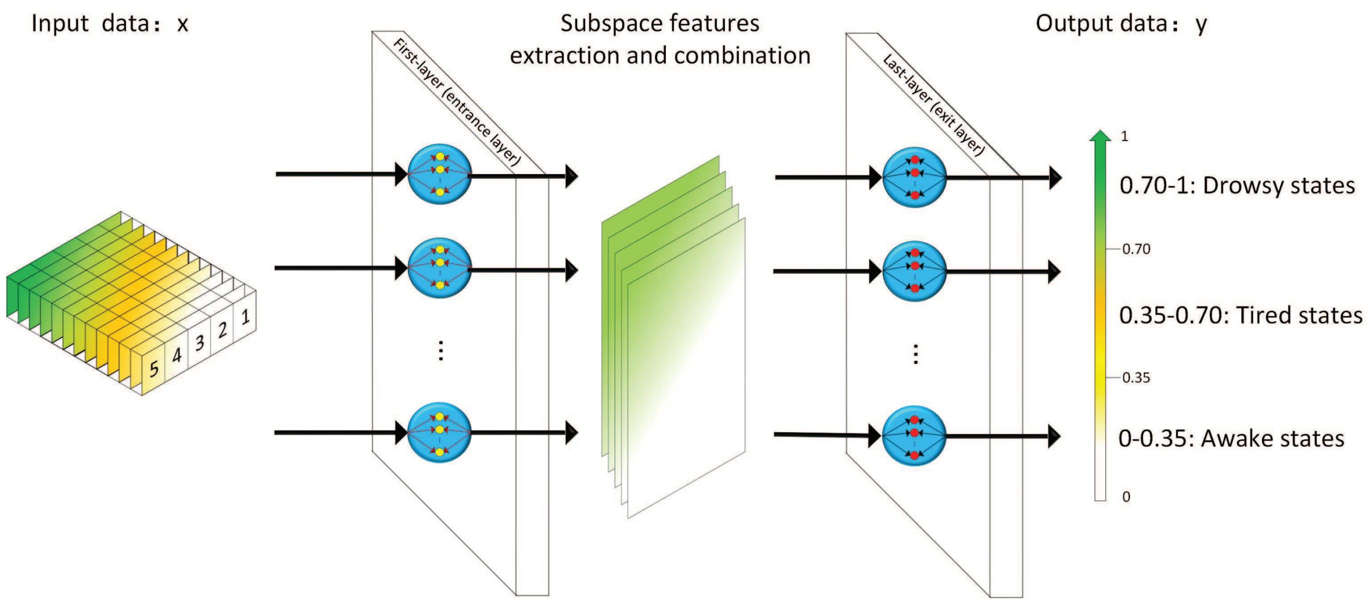


Fig. 1. The framework of the proposed method with forehead-EOG or EEG display the process of subspace features extraction and combination for the single modality.

TABLE I
SYMBOLS USED FOR THE PROPOSED METHOD.

Symbol	Property
R	the set of real numbers.
M	number of training samples.
$\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{M}$	x represents the input data and y represents the desired output data.
ω_i	the weight connecting the ith hidden node and the input nodes.
b_i	the bias of the ith hidden node.
β_i	the output weight between the ith hidden node and the output nodes.
$sum(e)$	the sum of all elements of the matrix residual error e.
$\hat{\omega}_p^i$	input weight of the ith subnetwork node in entrance layer. $\hat{\omega}_p^j \in \mathbb{R}^{d\times m}$
$\hat{\omega}_q^i$	input weight of the ith subnetwork node in exit layer. $\hat{\omega}_i^q \in \mathbb{R}^{d\times n}$
$\hat{b}_p^i$	bias of the ith subnetwork node in entrance layer. $\hat{b}_p^j \in \mathbb{R}$
(ω_{pi}^j, b_p^j)	the ith hidden node in the jth subnetwork node.
L	normalized function.
L^{-1}	reverse function of L.
$\mathbf{H}_p^j$	feature data generated by jth subnetwork nodes.
m	input data dimension.
n	output data dimension.
d	feature data dimension.
$\mathbf{e}_l$	the residual error of current network.
l	the numbers of subnetwork nodes.
S	sigmoid activation function.
η	the learning rate in the entrance layer.

orthogonal random. When the initial index $k = 1$, the initial subspace features of subnetwork node $\mathbf{H}_P^k$ are

$$
\begin{aligned}
&\mathbf{H}_p^k = S(\hat{\omega}_p^k, \hat{b}_p^k, \mathbf{x}) \\
&(\hat{\omega}_p^k)^T \cdot \hat{\omega}_p^k = \mathbf{I} \\
&(\hat{b}_p^k)^T \cdot \hat{b}_p^k = \mathbf{I}
\end{aligned} \tag{1}
$$

Step (2): For the exit layer, given the S activation function for any continuous desired outputs $\mathbf{y}$, the features of the subnetwork node $(\hat{\omega}_q^k, \hat{b}_q^k)$ are obtained by

$$
\begin{aligned}
&\hat{\omega}_q^k = S^{-1}\Big(L(\mathbf{y})\Big) \cdot \Big(S(\hat{\omega}_p^k, \hat{b}_p^k, \mathbf{x})\Big)^{-1} \\
&\hat{b}_q^k = \sqrt{mse\Big(\hat{\omega}_q \cdot S(\hat{\omega}_p^k, \hat{b}_p^k, \mathbf{x}) - S^{-1}\Big(L(\mathbf{y})\Big)\Big)}
\end{aligned} \tag{2}
$$

where $\mathbf{H}^{-1} = \mathbf{H}^T(\frac{U}{\mathbf{I}} + \mathbf{H}\mathbf{H}^T)^{-1}$, U represents a regularization value $(U > 0)$, $\hat{\omega}_q^k \in \mathbb{R}^{d\times n}$, and $\hat{b}_m^k \in \mathbb{R}$.

Step (3): Update $\mathbf{e}_k$, $\hat{\omega}_p^k$, and $\hat{b}_p^k$ as

$$
\begin{aligned}
&\mathbf{e}_k = \mathbf{y} - L^{-1} S(\mathbf{H}_p^k, \hat{\omega}_q^k, \hat{b}_q^k) \\
&\hat{\omega}_p^k = S^{-1}\Big(L(\mathbf{P}_{k-1} + \mathbf{H}_P^k)\Big) \cdot \mathbf{x}^{-1} + \eta \cdot \Big(S^{-1}\Big(L(\mathbf{P}_{k-1} + \mathbf{H}_P^k)\Big) \cdot \mathbf{x}^{-1}\Big) \\
&\hat{b}_p^k = \sqrt{mse(\hat{\omega}_p^k \cdot \mathbf{x} - \mathbf{P}_{k-1})}
\end{aligned} \tag{3}
$$

where $\mathbf{e}_k$ feedback the data $\mathbf{P}_k = S^{-1}(L(\mathbf{e}_k)) \cdot (\hat{\omega}_q^k)^{-1}, \mathbf{P}_0 = 0$, $\hat{\omega}_p^k \in \mathbb{R}^{m\times d}$, $\hat{b}_p^k \in \mathbb{R}$, the value of the learning rate η is 0.5.

Step (4): By setting $k = k+1$, we can obtain the k^{th} subspace features $(\hat{\omega}_p^k, \hat{b}_p^k)$ and the $(k+1)^{th}$ subspace features $(\hat{\omega}_p^{k+1}, \hat{b}_p^{k+1})$ as

$$
\begin{aligned}
&\mathbf{H}_p^k = S(\mathbf{x}, \hat{\omega}_p^k, \hat{b}_p^k) \\
&\mathbf{H}_p^{k+1} = S(\mathbf{x}, \hat{\omega}_p^{k+1}, \hat{b}_p^{k+1})
\end{aligned} \tag{4}
$$

Step (5): By repeating steps (2) to (4) $l-1$ times, we can obtain the final subspace features $\{\mathbf{H}_p^1, \cdots, \mathbf{H}_p^l\}$.

In general, the weight is obviously optimized after using the learning rate in Fig. 2. The green dots of the space α and red dots of the space β represent the original and updated data, respectively.

B. Feature Fusion

We use early fusion with max pooling for the feature fusion due to previous studies [49], [50] which indicate

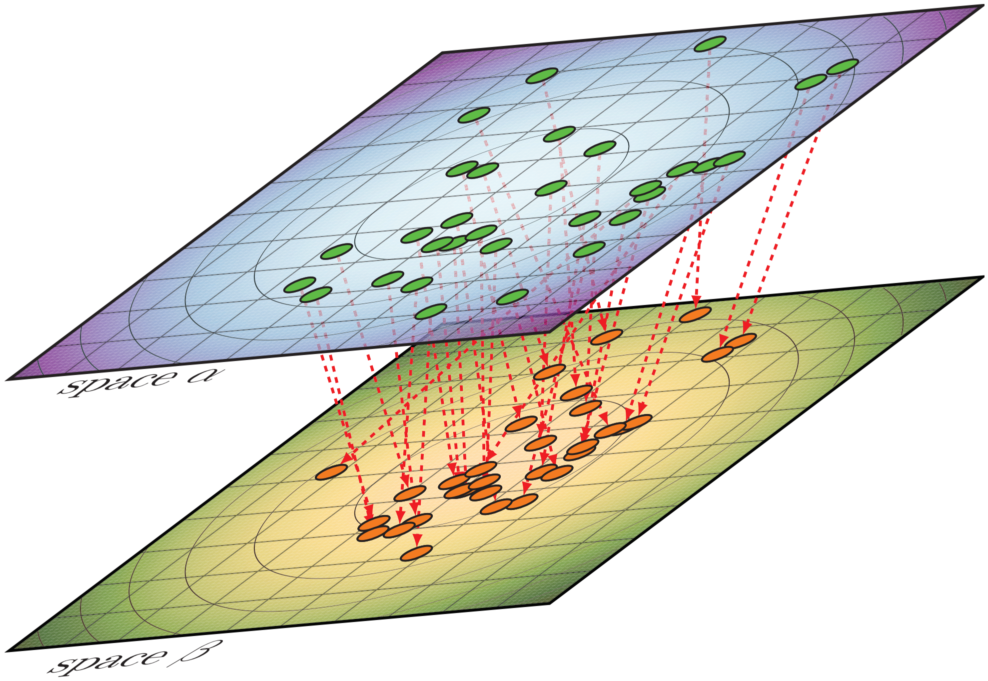


Fig. 2. The framework of the proposed method with the learning rate updated the weight of the entrance layer.

that an early fusion with kernel levelperformance is mostly robust and effective for the multilayer feature fusion. It is well known that in many popular convolutional neural network (CNN) models, such as Alexnet [51] and GoogleNet [52], max pooling was employed to reduce the deviation of the mean estimation that is caused by the convolution parameter error, which is widely used for reducing dimension and feature combination. For example, we define two sets of subspace features ($H^1 = H^1_1, H^1_2, \cdots, H^1_k$) and ($H^2 = H^2_1, H^2_2, \cdots, H^2_k$) that come from the entrance layer and the features $H^{1\oplus 2} = \max(H^1, H^2)$ fused by max pooling. The process of features fusion can be expressed as

$$\begin{aligned} H^{1\oplus 2} &= J(H^1, H^2) \\ H^{1\oplus 2\oplus 3} &= J\Big(J(H^1, H^2), H^3\Big) \\ &\vdots \\ H^{1\oplus 2\oplus\cdots\oplus n} &= J\Big(\cdots J\Big(J(H^1, H^2), H^3\Big), \cdots H^k\Big) \end{aligned} \tag{5}$$

where J is a combination operator.

C. Regression Model

We know that mixed neurons play an essential role in brain encoding and functions, and their subspace features can be expected to remove relevant factors of the brain. Meanwhile, generation of stable and complex behavior by the subspace features can be recast into the mapping space. Fig. 1 shows the details of the subspace features extraction of the proposed regression model for single modality. For multimodality, the process of the proposed method from data processing to regression analysis is illustrated in Fig. 3, which reflects the learning dimensions and structures that correspond with the biological evidence presented above. Here, the fused features represent the input data. The values of the output data in the range of '0-0.35', '0.36-0.70', and '0.71-1' represent awake, tired, and drowsy states, respectively. Furthermore, we proposed a theorem for regression model and specific content as:

Given the distinct N samples $\Big((X_i, t_i)^N_{i=1}, X_i \in \mathbb{R}^m, t_i \in \mathbb{R}^n\Big)$, S activation function and the arbitrary continuous desired output values t, the equation $\lim_{k\to\infty}\Big|\Big|t - L^{-1}\Big(S(\hat{\omega}^1_P \cdot X + \hat{b}^1_P)\Big)\cdot\hat{\beta}^1_P + \cdots + L^{-1}\Big(S(\hat{\omega}^k_P \cdot X + \hat{b}^k_P)\Big)\cdot\hat{\beta}^k_P\Big|\Big| \equiv 0$ holds when

$$\begin{aligned} \hat{\omega}^k_P &= S^{-1}\Big(L(e_{n-1})\Big)\cdot X^T(\frac{V}{I} + XX^T)^{-1}, \hat{\omega}^k_P \in \mathbb{R}^{m\times n} \\ \hat{b}^k_P &= \sum\frac{\Big(\hat{\omega}^k_P \cdot X - S^{-1}\Big(L(e_{m-1})\Big)\Big)}{N}, \hat{b}^k_P \in \mathbb{R} \\ S^{-1} &= -log(\frac{1}{x} - 1) \\ \beta^k_P &= \frac{\Big(e_{m-1}, L^{-1}\Big(q(\hat{\omega}^k_m \cdot X + \hat{b}^k_m)\Big)\Big)}{\Big|\Big|L^{-1}\Big(h(\hat{\omega}^k_m \cdot X + \hat{b}^k_m)\Big)\Big|\Big|}, \end{aligned} \tag{6}$$

where $X^T(\frac{V}{I} + XX^T)^{-1} = X^{-1}$ represents the Moore-Penrose generalized inverse of the training samples; S^{-1} is the inverse of activation function S; and L is the normalization function with the range of the data values $(0, 1]$; L^{-1} is the inverse of function L with the range of the data back to initial values.

D. Data Processing

1) Forehead EOG: According to different electrode placements, two EOG signals can be evaluated by forehead EOG

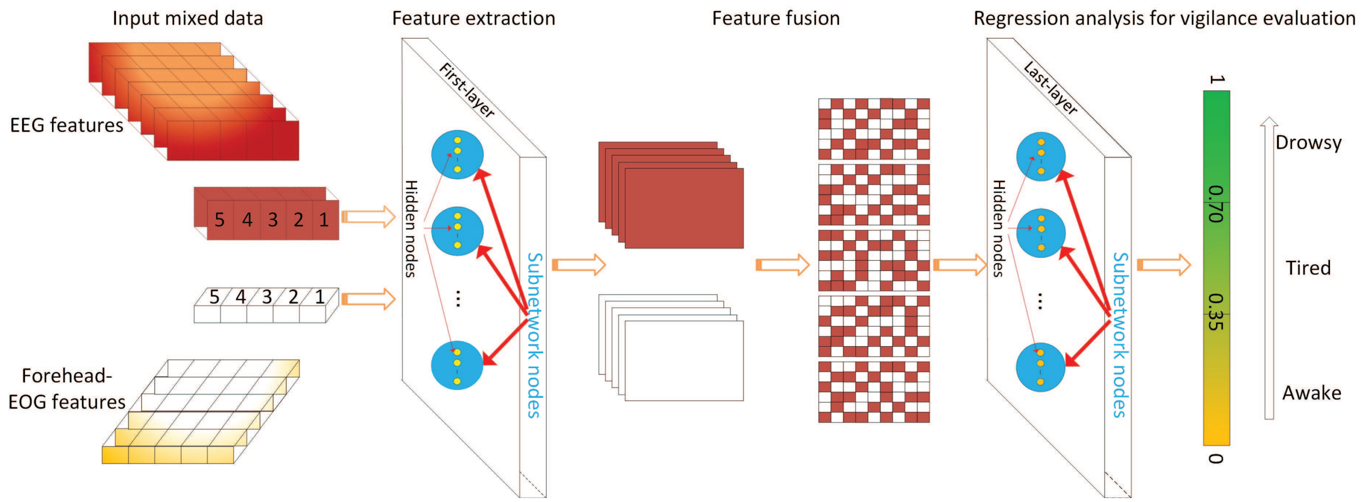


Fig. 3. The framework of the proposed method with the fused features display the process from data processing to regression analysis for multimodality.

and traditional EOG electrode placement. As seen in Fig. 4, two kinds of EOG features, namely vertical-EOG (VEO) and horizontal-EOG (HEO), were collected from electrodes No.1, and No.2 and electrodes No.3, and No.4, respectively. We used two separate approaches—independent component analysis (ICA) [53], [54] and the MINUS rule—to extract forehead-VEO (VEO_f-ICA) signals and forehead-HEO (HEO_f-MINUS) signals. Here, 'forehead' is denoted by the subscript 'f'. Zhang *et al.* [18] also reported that forehead-EOG signals, similar to the traditional EOG signals, contain crucial eye movements, such as blink, fixation, and saccade components.

We used the valid wavelet transform approach, Mexican hat wavelet, to detect eye movements and the formula as

$$\psi(t) = \frac{2}{\sqrt{3\sigma}\pi^{\frac{1}{4}}}\left(1-(\frac{t}{\sigma})^2\right)e^{\frac{-t^2}{2\sigma^2}} \tag{7}$$

where σ represents the standard deviation. We encoded the process of the wavelet transform, and the negative peak and positive peak were encoded as '0' and '1', respectively. Simultaneously, blink and saccade features can be represented by '01' or '10' and '010', respectively. Finally, the total of 36 EOG features extracted by the detected eye movements are shown in Table II.

2) Forehead EEG: Based on previous studies, the eye movements and the blink artifacts are included in the EEG recording, which also contains crucial information for vigilance estimation. To extract EOG and EEG signals that have been recorded from the No.3 to the No.4 forehead electrodes, we use the fast independent component analysis (FASTICA) approach [55] to separate EOG and EEG signals. Then, the forehead EEG components are reconstructed and encoded as an input matrix X by ICA algorithm; we can obtain the un-mixing matrix W after decomposition and, finally, the forehead EEG signals can be derived through the formulas as

$$\begin{aligned} X &= [ch\_1; ch\_2; -ch\_3; ch\_4] \\ Z &= Y * X \\ \tilde{X} &= Y^{-1} * \tilde{Z} \end{aligned} \tag{8}$$

where Nos. 1-4 columns of matrix X represent No.1, No.2, No.3, and No.4 channels, respectively; $\tilde{Z}$ is a matrix for activation waveforms Z and its rows consist of EOG components that have been set to zero.

3) Temporal and Posterior EEG: In additional, Shi and Lu [56] reported that the EEG signals of temporal and posterior sites have important information for vigilance estimation. To reduce noise and artifacts, the raw EEG signals are evaluated by a band-pass filter, with a frequency of 1 to 50 Hz in pre-processing. They also indicate that Differential Entropy (DE) has a promising capability for vigilance estimation, from low to high frequency energy; the formula is for this is given by

$$h(X) = -\int_X f(x)log\Big(f(x)dx\Big) \tag{9}$$

If the random time series X follows the Gauss distribution $N(\mu, \delta)$, the DE features are defined as

$$h(X) = -\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\delta^2}}e^{-\frac{(x-\mu)^2}{2\delta^2}}\,log\Big(\frac{1}{\sqrt{2\pi\delta^2}}e^{-\frac{(x-\mu)^2}{2\delta^2}}\Big)dx = log(2\pi\delta^2) \tag{10}$$

We used two methods, including five frequency bands (1-50 Hz) with a 2 Hz frequency resolution (2Hz) and five frequency bands (5Bands), to extract differential entropy (DE) and power spectral density (PSD) features from EEG signals. The five frequency bands were as follows: delta (1-4 Hz), theta (4-8 Hz), alpha (8-14 Hz), beta (14-31 Hz), and gamma (31-50 Hz). Then, four features, namely PSD-MA, PSD-LDS, DE-MA, and DE-LDS, were extracted by separation methods of moving average (MA) and linear dynamic system (LDS) filtering. This feature is shown in

TABLE II
THE FOREHEAD-EOG SIGNALS EXTRACTED FROM 36 EYE MOVEMENTS

Group	Features
Saccade	maximum/minimum/mean of saccade rate/saccade amplitude, maximum/mean of saccade rate variance/saccade amplitude variance, power/mean power of saccade amplitude, saccade number
Blink	maximum/mean of blink rate variance/amplitude variance maximum/minimum/mean of blink amplitude, power/mean power of blink amplitude, maximum/mean/sum of blink rate, blink number
Fixation	maximum/mean of blink duration variance/saccade duration variance, maximum/minimum/mean of blink duration/saccade duration

Table III. Moreover, we simultaneously recorded two other EEG signals from the human brain, including the posterior site (12-channels: CPZ, CP1, CP2, PZ, P1, P2, POZ, PO3, PO4, OZ, O1, and O2) and the temporal site (6-channels: T7, T8, FT7, FT8, TP7, and TP8) by 10-20 international electrode system shown in Fig. 5.

Finally, the bandpass filtering method with a frequency range of 1 to 50 Hz is used to first remove the effects of myoelectricity in the original signal and then noise and artifacts that have a significant impact on the data.

E. Vigilance Estimation

The annotation of the PERCLOS is used to judge the drowsy state by the percentage of eye closure, which is one of the most widely accepted vigilance indices in literature. Conventional facial video-based technology can be easily influenced by environmental changes,such as massive occlusion and various illuminations [57]–[59]. Here, we adopt an automatic continuous vigilance technique based on the eye-tracking-glasses to calculate how precisely does the PERCLOS reflect eye movements, including blink, saccade, and fixation components. This has been proved by Gao *et al.* [19] in both real and laboratory environments. The SMI-ETG eye tracking glasses[2] are shown in Fig. 6. The formula for PERCLOS is

$$PERCLOS = \frac{blink + CLOS}{blink + saccade + fixation + CLOS} \tag{11}$$

where 'CLOS' represents the duration of the eye closures—usually considered as the eyelids covering the pupil by over 80%.

The exit layer with the data is obtained by subspace feature extraction and fusion. All regression models will be introduced in the next section.We use two critical indices, i.e., root-mean-square error (RMSE) and Pearson product-moment correlation coefficient (COR) to finally evaluate levels of vigilance. RMSE is frequently defined as the squared error between the observed and predicted values and the formula is

$$RMSE(x, y) = \sqrt{\frac{\sum_{t=1}^{n}(x_t - y_t)^2}{n}} \tag{12}$$

where $x = (x_1, x_2, ..., x_n)^T$ represent the observed values and $y = (y_1, y_2, ..., y_n)^T$ the predicted values by the regression model.

[2]http://eyetracking-glasses.com/

We use the COR to overcome the disadvantage that RMSE cannot obtain the relationship that is established between the observed and the predicted values. The COR values in the range from '-1' to '+1' describe the linear relationship between the observed and predicted values, where '-1' represents the most possible disagreement, '0' represents no relationship, and '+1' represents the most possible agreement. In general, the higher the accuracy of the regression analysis with the lower RMSE, the higher the COR. The formula of COR is

$$COR(x, y) = \frac{\sum_{t=1}^{n}(x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^{n}(x_t - \bar{x})^2 \sum_{t=1}^{n}(y_t - \bar{y})^2}} \tag{13}$$

where $\bar{x}$ represents the mean of x and $\bar{y}$ represents the mean of y.

III. EXPERIMENTAL SETUP

A. Environment Setting

The forehead-EOG and EEG dataset (SEED-VIG) was collected by Zheng *et al.* [22]. There are 23 subjects (11 males and 12 females with an average age of 23.3 years) who participated in the experiments. All participants were without the influence of any kind of drugs such as caffeine, alcohol, and tobacco, etc. and possessed normal hearing and self-reported normal or corrected-to-normal vision. As we know, Ferrara *et al.* [60] reported that humans are completely sleepy approximately at 1:30 pm after lunch, when their fatigue can quickly reach the peak. The experiments were performed at that time and lasted about two hours without alertness in the simulated environments. As seen in Fig. 7, there is an LCD screen that comprises of a four-lane highway, simulated as a real environment in front of the experimental vehicle and the movements of the vehicle without any dynamical system; for instance, the engine controls its movement by software, the gas pedal, and the steering wheel. The simulated environment can be updated in real time; the subjects of the experiments are sleeping during driving, without any warning feedback. Both forehead-EOG and EEG raw signals were collected from ESI NeuroScan System using a 1000 HZ sampling frequency.

B. Compared Methods

In this section, we tested our methods with the forehead-EOG and EEG datasets that were collected by the simulated

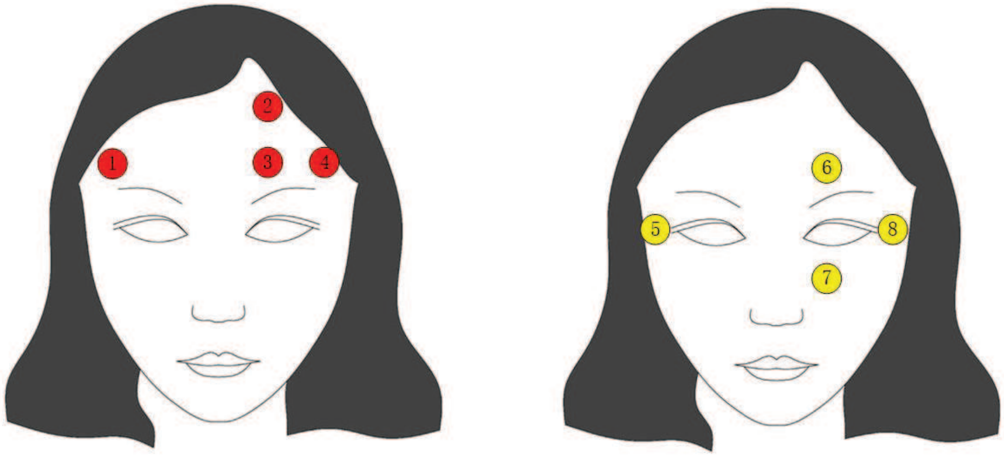


Fig. 4. Different electrode placements for extraction of the forehead-EOG and traditional signals. The four red dots and yellow dots represent the electrode placement of the forehead EOG and the traditional EOG, respectively.

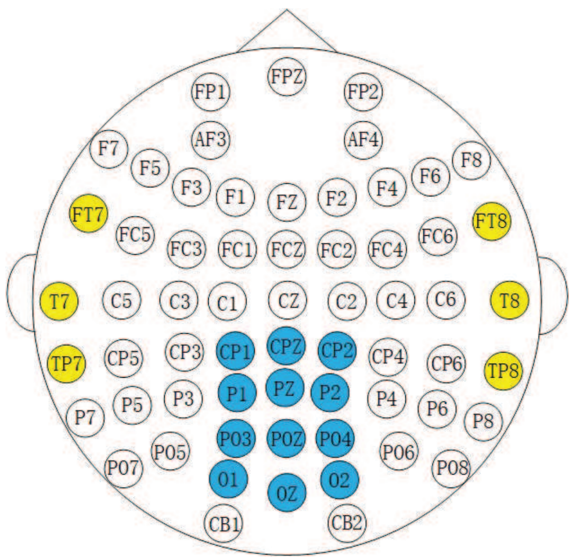


Fig. 5. EEG signals collected from different sites by different electrode placements. The yellow dots represent 6-channel EEG signals of the temporal site and the blue dots represent 12-channel EEG signals of the posterior site.

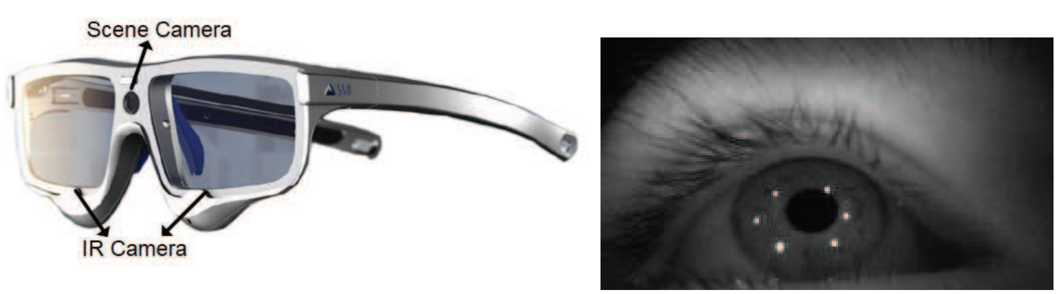


Fig. 6. Eye movements recorded by the SMI eye tracking glasses.

driving system. There are 23 experiments in total, and each experiment comprises 885 samples of forehead-EOG and EEG features. For evaluation levels of vigilance, we separate the entire data from one experiment into five ses-

Fig. 7. The experimental LCD screen of the simulated driving system.

sions and evaluate the performance with a five-fold cross-validation and obtained the root-mean-square error (RMSE) and correlation coefficient (COR), which are used as the final evaluations. These regression models are listed below: 1) ICA, ICAV-MINH, MINUS, ELM [61], B-ELM [62], and the proposed method with or without the learning rate using forehead-EOG for single modality; 2) ICA, ELM, B-ELM, and the proposed method with or without the learning rate using EEG for single modality; 3) SVR [63], CCRF [64], CCNF [65], S-LSTM, F-LSTM [66], ELM, B-ELM, Autoencoder-ELM (Auto-ELM) [67], and the proposed method with or without the learning rate using the fused features for multimodality. We introduce two different long short-term memory (LSTM) encoders, in which one encodes EOG and EEG into a compact feature vector by stacked LSTM layers (S-LSTM) and the another EOG and EEG into one feature vector (F-LSTM). Simultaneously, we used the autoencoder model to reduce the dimensionality of the input data and combined it with the ELM regression network (Auto-ELM) for vigilance estimation.

We used Matlab 2017b with 32GB memory to test our algorithm and compared methods for single modality and multimodality. The code that could be downloaded publicly from the internet[3], and the valuation of parameters can be adjusted in every step experiment.Here, the learning rate η for the entrance layer was tuned in the range of $(0,1)$ and we chose a value of 0.8. The optimal values of the regularization parameter V were tuned in the range of $[2^{-9}, 2^{-8}, \ldots, 2^{9}]$, and $[2^{-9}, 2^{-8}, \ldots, 2^{5}]$ for single modality and multimodality, respectively. The range of regularization parameters α_k and β_k to train the CCRF and CCNF were $10^{[0,1,2]}$ and $10^{[-3,-2,-1,0]}$, respectively. The number of vertex features for five-fold cross-validation are $K_1 = [10, 20, 30]$. If i^{th} and j^{th} nodesare neighbors, then $K_2 = 1$ and $S^k = 1$; otherwise, $K_2 = 0$ and $S^k = 0$.

IV. Experimental Evaluation

A. Using EOG for Single Modality

The forehead-EOG feature include HEO$_f$-ICA, VEO$_f$-ICA, HEO$_f$-MINUS and VEO$_f$-MINUS extracted from ICA and

[3]http://www.yiminyang.com/

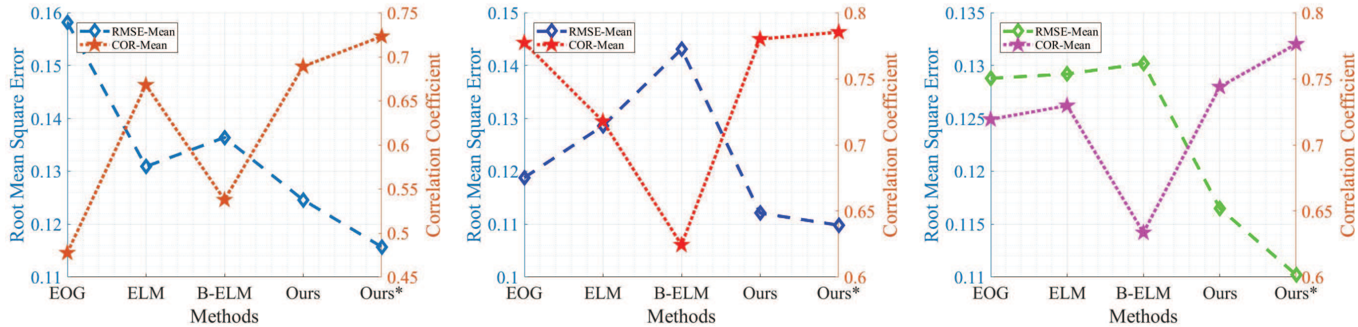
TABLE III
The detailed information of forehead-EOG and different sites EEG from one of twenty three subjects. Every subjects' feature from the same site has the same number of samples and dimensions. We use $x \times y$ to describe the number of samples and dimensions, i.e., 885 and 36 represent samples and dimensions in 885×36, respectively.

	EOG-ICA	EOG-MINUS	EOG-ICAV-MINH	
Forehead-EoG	885×36	885×36	885×36	
	DE-LDS	DE-MA	PSD-LDS	PSD-MA
Forehead-EEG-2HZ	885×100	885×100	885×100	885×100
Forehead-EEG-5BANDS	885×20	885×20	885×20	885×20
Temporal-EEG-2HZ	885×150	885×150	885×150	885×150
Temporal-EEG-5BANDS	885×30	885×30	885×30	885×30
Posterior-EEG-2HZ	885×275	885×275	885×275	885×275
Posterior-EEG-5BANDS	885×55	885×55	885×55	885×55

TABLE IV
The mean RMSE, COR, and their standard deviations (STD) of compared methods use EOG. The best results are bolded. Ours* – ours use the learning rate.

Methods	RMSE-Mean	RMSE-STD	COR-Mean	COR-STD
EOG-ICA	0.1582	0.0844	0.4774	0.5381
ELM	0.1309	0.0486	0.6680	0.1957
B-ELM	0.1364	0.0683	0.5376	0.1906
Ours	0.1246	0.0540	0.6890	0.2041
Ours*	**0.1157**	**0.0445**	**0.7231**	**0.1832**
EOG-ICAV-MINH	0.1188	0.0391	0.7773	0.2352
ELM	0.1286	0.0557	0.7178	0.2100
B-ELM	0.1431	0.0584	0.6243	0.1904
Ours	0.1121	0.0540	0.7801	0.2041
Ours*	**0.1098**	**0.0402**	**0.7853**	**0.1875**
EOG-MINUS	0.1288	0.0588	0.7193	0.3492
ELM	0.1292	0.0489	0.7297	0.2225
B-ELM	0.1302	0.0596	0.6336	0.1984
Ours	0.1165	0.0508	0.7439	0.1819
Ours*	**0.1102**	**0.0417**	**0.7763**	**0.1748**

MINUS separation approaches. We tested the comparative methods including ICA, ICAV-MINH, MINUS, ELM, B-ELM, and the proposed algorithm with or without the learning rate using forehead-EOG. The experimental results of the mean RMSE, the mean COR, and their standard deviations (STD) are shown in Table IV. Simultaneously, we evaluated the statistical significance using a one-way analysis of variance (ANOVA). The mean RMSE/COR of the EOG-ICA, ICAV-MINH, and MINUS were seen to be as 0.16/0.48, 0.12/0.78, and 0.13/0.72,respectively; the ICAV-MINH-based method can obtain a better result due to its easy detectability of blink and saccade components from



(a) EOG features collected by ICA. (b) EOG features collected by ICAV-MINH. (c) EOG features collected by MINUS.

Fig. 8. Different simulations for single modality using EOG that was collected by different separation approaches. Here, 'ICAV' represents VEO features collected by ICA. 'MINH' represents HEO features collected by MINUS.

TABLE V
THE EXPERIMENTAL RESULTS WITH DE AND PSD FEATURES FROM DIFFERENT SITES. THE BEST RESULTS ARE BOLDED. OURS* – OURS USE THE LEARNING RATE.

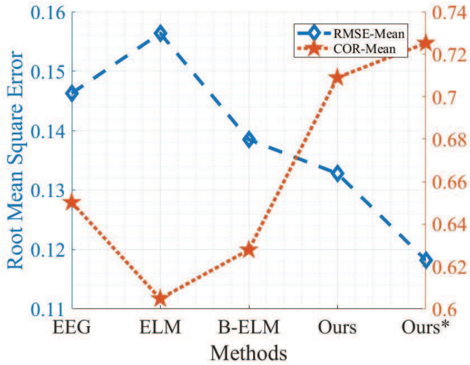
Features	Parameters	DE-LDS	DE-MA	PSD-LDS	PSD-MA
		Ours/Ours*	Ours/Ours*	Ours/Ours*	Ours/Ours*
Forehead-EEG-2Hz	RMSE-Mean	0.1328/**0.1182**	0.1427/0.1263	0.1452/0.1369	0.1501/0.1418
	RMSE-STD	0.0384/0.0372	0.0373/**0.0369**	0.0409/0.0446	0.0414/0.0411
	COR-Mean	0.7090/**0.7253**	0.6666/0.7020	0.6351/0.6863	0.6100/0.6499
	COR-STD	0.1769/0.1750	0.1873/**0.1718**	0.1954/0.1863	0.2066/0.1863
Forehead-EEG-5Bands	RMSE-Mean	0.1363/**0.1181**	0.1415/0.1245	0.1528/0.1392	0.1539/0.1435
	RMSE-STD	0.0386/0.0398	**0.0385**/0.0403	0.0455/0.0477	0.0435/0.0431
	COR-Mean	0.7041/**0.7355**	0.6816/0.7095	0.6269/0.6627	0.6094/0.6226
	COR-STD	0.1781/**0.1761**	0.1895/0.1835	0.2041/0.2033	0.2867/0.2310
Temporal-EEG-2Hz	RMSE-Mean	0.1251/**0.1157**	0.1306/0.1190	0.1408/0.1264	0.1442/0.1339
	RMSE-STD	0.0341/**0.0336**	0.0342/0.0339	0.0395/0.0362	0.0359/0.0357
	COR-Mean	0.7041/**0.7283**	0.6791/0.7051	0.6559/0.6922	0.6465/0.6668
	COR-STD	0.1806/0.1731	0.1880/0.1759	0.1902/**0.1703**	0.1928/0.1768
Temporal-EEG-5Bands	RMSE-Mean	0.1361/**0.1318**	0.1399/0.1349	0.1407/0.1388	0.1485/0.1451
	RMSE-STD	0.0590/0.0582	0.0514/0.0526	**0.0497**/0.0520	0.0536/0.0549
	COR-Mean	0.6651/**0.6796**	0.6419/0.6585	0.6569/0.6644	0.6144/0.6250
	COR-STD	0.2141/0.2077	0.2165/0.2275	0.2232/**0.2058**	0.2261/0.2265
Posterior-EEG-2Hz	RMSE-Mean	0.1251/**0.1175**	0.1321/0.1239	0.1498/0.1399	0.1513/0.1456
	RMSE-STD	0.0452/0.0420	0.0416/0.0414	0.0398/**0.0394**	0.0411/0.0441
	COR-Mean	0.7167/**0.7201**	0.6743/0.6823	0.6487/0.6676	0.6390/0.6591
	COR-STD	0.1755/**0.1706**	0.1957/0.2076	0.2099/0.1907	0.2287/0.2178
Posterior-EEG-5Bands	RMSE-Mean	0.1270/**0.1240**	0.1304/0.1262	0.1399/0.1366	0.1457/0.1430
	RMSE-STD	**0.0444**/0.0485	0.0447/0.0467	0.0488/0.0457	0.0476/0.0479
	COR-Mean	0.7215/**0.7414**	0.6999/0.7150	0.6676/0.7081	0.6444/0.6694
	COR-STD	0.1863/**0.1791**	0.1934/0.1926	0.2032/0.1852	0.2108/0.1895

forehead-EOG signal.

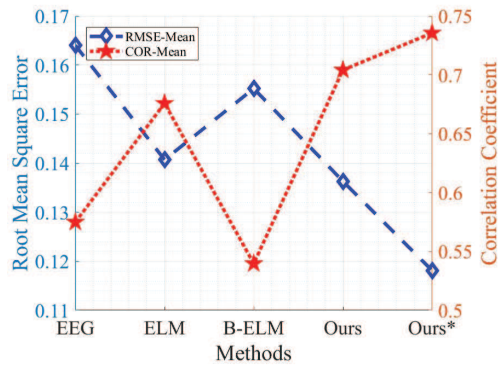
Compared to the results of other regression models, the profit of our strategies using forehead-EOG feature is apparent in Fig. 8. The Extreme Learning Machine (ELM)-based method for vigilance estimation works reasonably, as it can achieve a good RMSE/COR of 0.1309/0.6680, 0.1286/0.7178, and 0.1292/0.7297, respectively. The results of the proposed method without the learning rate are 0.1246/0.6890, 0.1121/0.7801 and 0.1165/0.7439, respectively. After using the learning rate, the results improved to 0.1157/0.7231, 0.1098/0.7853, and 0.1102/0.7763, respectively. Our approaches with the learning rate obtained the lowest RMSE and the highest COR and the mean results significantly improved to 0.1119/0.7616 ($p<0.01$/$p<0.01$, ANOVA), which demonstrated that it has better performance in detecting saccade, blink, and fixation component compared to other state-of-the-art techniques.

B. Using EEG for Single Modality

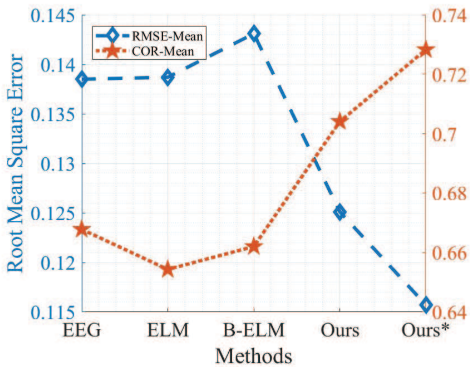
As mentioned above, the forehead-EEG, temporal-EEG, and posterior-EEG were extracted by 4-channels of the forehead site, 6-channels of the temporal site, and 12-channels of the posterior site, respectively. According to the entire experimental results in Table V, we found that the DE-LDS feature has a promising effect. In addition to the results of the EEG setup from different electrode placements, we used two different methods, including the 2 Hz frequency resolution and the five frequency bands,



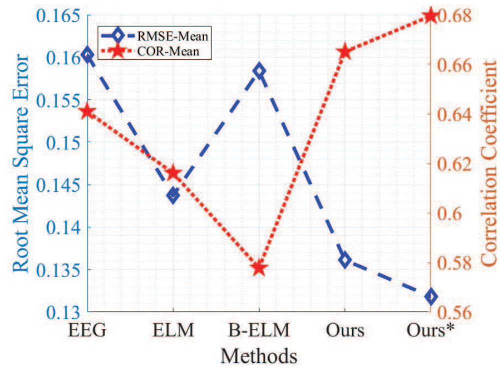
(a) Forehead-EEG features collected by 2Hz.



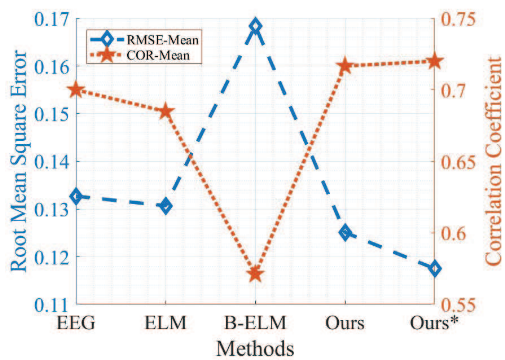
(b) Forehead-EEG features collected by 5Bands.



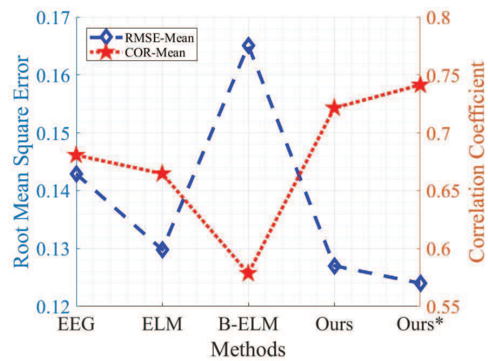
(c) Temporal-EEG features collected by 2Hz.



(d) Temporal-EEG features collected by 5Bands.



(e) Posterior-EEG features collected by 2Hz.



(f) Posterior-EEG features collected by 5Bands.

Fig. 9. Different simulations for single modality using EEG collected by different sites of the brain. Ours* – ours use the learning rate.

to analyze the data collected from the same position. The experimental results that were computed using the 2 Hz frequency resolution were better than those five frequency bands with the same feature. This is also consistent with the previous studies.

We utilize comparison methods of EEG-ICA, ELM, B-ELM, and the proposed method with or without the learning rate using the DE feature extracted by two ways from different sites of the human brain to estimate vigilance. The experimental results are shown in Table VI, and it was found that the EEG-ICA approach has well performed consistently with previous conclusions [22]. It was noted that the proposed approach has an outstanding performance, it had obtained a reduced RMSE and increased COR of 3%/10%. The proposed method with the learning rate obtained the lowest RMSE and highest COR of 0.1175/0.7414 ($p<0.01$/$p<0.01$, ANOVA) for single modality; its effectiveness has been verified in Fig. 9, which indicates that EEG signal included the critical information for estimating the level of vigilance.

C. Using the Feature Fusion for Multimodality

In this part, we used the feature fusion of forehead-EOG and EEG to evaluate the levels of vigilance for multimodality. According to the complementary characteristics of forehead EOG and EEG, combination of different sites of EEG features (forehead-EEG, temporal-EEG, and posterior-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2018.2889223, IEEE Transactions on Cognitive and Developmental Systems

TABLE VI
THE MEAN RMSE, COR, AND THEIR STANDARD DEVIATIONS (STD) OF COMPARED METHODS WITH DE FEATURES. THE BEST RESULTS ARE BOLDED. OURS* – OURS USE THE LEARNING RATE.

Features	Methods	RMSE-Mean	RMSE-STD	COR-Mean	COR-STD
Forehead-2Hz	EEG-ICA	0.1463	0.0383	0.6502	0.2116
	ELM	0.1564	0.0505	0.6048	0.1866
	B-ELM	0.1385	0.0434	0.6278	0.1894
	Ours	0.1328	0.0384	0.7090	0.1769
	Ours*	**0.1182**	**0.0372**	**0.7253**	**0.1750**
Forehead-5Bands	EEG-ICA	0.1640	0.0483	0.5749	0.2463
	ELM	0.1407	**0.0370**	0.6758	0.2317
	B-ELM	0.1552	0.0469	0.5397	0.1814
	Ours	0.1363	0.0386	0.7041	0.1781
	Ours*	**0.1181**	0.0398	**0.7355**	**0.1761**
Temporal-2Hz	EEG-ICA	0.1385	0.0343	0.6678	0.2349
	ELM	0.1387	0.0535	0.6543	0.2372
	B-ELM	0.1431	0.0584	0.6620	0.1904
	Ours	0.1251	0.0341	0.7041	0.1806
	Ours*	**0.1157**	**0.0336**	**0.7283**	**0.1731**
Temporal-5Bands	EEG-ICA	0.1603	0.0722	0.6410	0.1745
	ELM	0.1437	0.0591	0.6161	0.2100
	B-ELM	0.1584	0.0636	0.5777	**0.1665**
	Ours	0.1361	0.0590	0.6651	0.2141
	Ours*	**0.1318**	**0.0582**	**0.6796**	0.2077
Posterior-2Hz	EEG-ICA	0.1327	**0.0303**	0.7001	0.2250
	ELM	0.1307	0.0429	0.6849	0.2682
	B-ELM	0.1684	0.0514	0.5711	0.1968
	Ours	0.1251	0.0452	0.7167	0.1755
	Ours*	**0.1175**	0.0420	**0.7201**	**0.1706**
Posterior-5Bands	EEG-ICA	0.1429	**0.0393**	0.6807	0.2129
	ELM	0.1298	0.0438	0.6647	0.3580
	B-ELM	0.1651	0.0532	0.5785	0.1988
	Ours	0.1270	0.0444	0.7215	0.1863
	Ours*	**0.1240**	0.0485	**0.7414**	**0.1791**

TABLE VII
THE PERFORMANCE OF THE PROPOSED METHOD USING DIFFERENT LEARNING RATES FOR MULTIMODALITY.

Learning rate η	RMSE-Mean	COR-Mean
without η	0.09	0.85
$\eta = 0.10$	0.08	0.86
$\eta = 0.20$	0.09	0.86
$\eta = 0.50$	0.09	0.86
$\eta = 0.80$	0.08	0.86

TABLE VIII
THE MEAN RMSE AND COR OF COMPARED METHODS FOR MULTIMODALITY USING THE FEATURE FUSION. THE BEST RESULTS ARE BOLDED. OURS*-MAX POOLING – OURS USE THE LEARNING RATE WITH MAX POOLING.

Methods	RMSE-Mean	COR-Mean
SVR	0.10	0.83
CCRF	0.10	0.84
CCNF	0.09	0.85
S-LSTM	**0.08**	0.83
F-LSTM	**0.08**	0.84
ELM	0.11	0.78
B-ELM	0.12	0.70
Autoencoder-ELM	0.10	0.82
Ours - max pooling	0.09	0.85
Ours* - max pooling	**0.08**	**0.86**

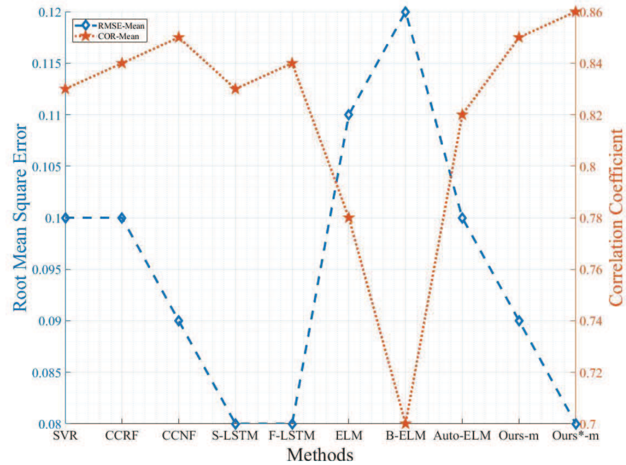


Fig. 10. Different simulations for the multimodality using the feature fusion. Ours-m – ours with max pooling. Ours*-m – ours use the learning rate with max pooling.

EEG) and forehead-EOG features for vigilance estimation, our experimental results showed the forehead-EOG combination of posterior-EEG has a promising effect with the lowest RMSE and the highest COR. It thus explains that the posterior-EEG has extra crucial information for vigilance estimation than the forehead-EEG. It was also observed that the forehead-EEG just uses 4-channel electrodes, which is much lesser than the posterior-EEG extracted by 12-channel electrodes. This makes the forehead electrode placement easier to commercialize with the low cost.

The different learning rates are applied to the proposed method, and the results are displayed in Table VII. The performance of the proposed multimodality regression model has significantly improved after using learning rate in the entrance layer and choosing an optimal result. Table VIII displays the results of the comparison methods of SVR, CCRF, CCNF, S-LSTM, F-LSTM, ELM, B-ELM, Auto-ELM, and the proposed method with or without the learning rate. We can observe that the CCRF and CCNF methods with temporal dependency obviously improved the mean RMSE/COR to 0.10/0.84, and 0.09/0.85, respectively. Although the LSTM recurrent neural networks (RNN) achieves a good RMSE, the ordinary performance of the COR reduces its practical effect significantly. Simultaneously, we observe the fact that an ELM-based model combined with the autoencoder layer, which reduced the dimensionality of the input data has improved the performance considerably, which is a promising future. Furthermore, the benefit of the multimodality is evident as forehead-EOG, and EEG signals have the characteristics of the complementarity [21] of the EOG and the EEG. The effectiveness and efficiency of the proposed method with the DE feature can be seen in Fig. 10. It was also observed that the proposed approach with a optimal value of the learning rate ($\eta = 0.8$) achieved the lowest RMSE and highest COR of 0.08/0.86 ($p<0.01/p<0.01$, ANOVA), which outperforms other state-of-the-art techniques for multimodality.

V. Conclusion

In this paper, we proposed a double-layered neural network with subnetwork nodes (DNNSN) for vigilance estimation, including using forehead-EOG and EEG for single modality and the feature fusion for multimodality. When the learning rate was applied to the input layer, the mean RMSE/COR of the proposed method for single modality using forehead-EOG and EEG features were 0.11/0.79, and 0.12/0.74, respectively, while that for multimodality utilizing the feature fusion was 0.08/0.86. As we know, EOG has two advantages of easy setup and a high signal-to-noise ratio, but it is easily influenced by video device and external environment. Likewise, although EEG can record neurophysiological signals of our brain that is correlated with alertness directly, it has a low signal-to-noise ratio. Multimodality employs the complementary advantages of the mixed signals to estimate the levels of vigilance and proves the correctness through experimental results. And then, the experimental results are improved significantly when the proposed method combined with the learning rate was used. Furthermore, the proposed multimodality algorithm achieve a lower standard deviation than other two single modalities, which proves the multimodality can also improve the robustness of the vigilance detection model. In general, we demonstrated the feasibility and efficiency of the proposed method for vigilance estimation using the feature fusion.

References

[1] J. S. Warm, "The psychophysics of vigilance," in *Proceedings of the Human Factors Society Annual Meeting*, vol. 24, pp. 605–605, Sage Publications Sage CA: Los Angeles, CA, 1980.

[2] B. D. Taylor and R. L. Goldingay, "Cutting the carnage: wildlife usage of road culverts in north-eastern New South Wales," *Wildlife Research*, vol. 30, no. 5, pp. 529–537, 2003.

[3] M. R. Rosekind, E. L. Co, K. B. Gregory, and D. L. Miller, "Crew factors in flight operations. 8; a survey of fatigue factors in corporate/executive a viation operations," *Nasa Ames Research Center Moffett Field Ca Technical Memorandum*, 2000.

[4] G. Zhang, K. K. Yau, and G. Chen, "Risk factors associated with traffic violations and accident severity in China," *Accident Analysis & Prevention*, vol. 59, pp. 18–25, 2013.

[5] G. Zhang, K. K. Yau, and X. Gong, "Traffic violations in guangdong province of China: speeding and drunk driving," *Accident Analysis & Prevention*, vol. 64, pp. 30–40, 2014.

[6] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 596–614, 2011.

[7] A. L. Loomis, E. N. Harvey, and G. A. Hobart, "Cerebral states during sleep, as studied by human brain potentials.," *Journal of Experimental Psychology*, vol. 21, no. 2, p. 127, 1937.

[8] M. Matousek and I. Petersén, "A method for assessing alertness fluctuations from EEG spectra," *Electroencephalography and Clinical Neurophysiology*, vol. 55, no. 1, pp. 108–113, 1983.

[9] H. Merica and J.-M. Gaillard, "The EEG of the sleep onset period in insomnia: a discriminant analysis," *Physiology & Behavior*, vol. 52, no. 2, pp. 199–204, 1992.

[10] S. Makeig and M. Inlow, "Lapse in alertness: coherence of fluctuations in performance and EEG spectrum," *Electroencephalography and Clinical Neurophysiology*, vol. 86, no. 1, pp. 23–35, 1993.

[11] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, "Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 121–131, 2011.

[12] N. Zhang, W.-L. Zheng, W. Liu, and B.-L. Lu, "Continuous vigilance estimation using lstm neural networks," in *Int. Conf. on Neural Information Processing (ICONIP)*, pp. 530–537, Springer, 2016.

[13] Y. Yang, Q. J. Wu, W.-L. Zheng, and B.-L. Lu, "EEG-based emotion recognition using hierarchical network with subnetwork nodes," *IEEE Trans. Cogn. Dev. Syst.*, 2017.

[14] R. Chai, S. H. Ling, P. P. San, G. R. Naik, T. N. Nguyen, Y. Tran, A. Craig, and H. T. Nguyen, "Improving EEG-based driver fatigue classification using sparse-deep belief networks," *Frontiers in Neuroscience*, vol. 11, p. 103, 2017.

[15] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea, and M. E. Lopez, "Real-time system for monitoring driver vigilance," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 63–77, 2006.

[16] P. S. Peopie, "Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management," *Drivers*, 1998.

[17] S. Hu and G. Zheng, "Driver drowsiness detection with eyelid related parameters by support vector machine," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7651–7658, 2009.

[18] Y.-F. Zhang, X.-Y. Gao, J.-Y. Zhu, W.-L. Zheng, and B.-L. Lu, "A novel approach to driving fatigue detection using forehead EOG," in *7th Int. IEEE/EMBS Conf. on Neural Engineering*, pp. 707–710, 2015.

[19] X.-Y. Gao, Y.-F. Zhang, W.-L. Zheng, and B.-L. Lu, "Evaluating driving fatigue detection algorithms using eye tracking glasses," in *7th Int. IEEE/EMBS Conf. on Neural Engineering*, pp. 767–770, 2015.

[20] J.-X. Ma, L.-C. Shi, and B.-L. Lu, "An EOG-based vigilance estimation method applied for driver fatigue detection," *Neuroscience and Biomedical Engineering*, vol. 2, no. 1, pp. 41–51, 2014.

[21] X.-Q. Huo, W.-L. Zheng, and B.-L. Lu, "Driving fatigue detection with fusion of EEG and forehead EOG," in *Int. Joint Conf. on Neural Networks*, pp. 897–904, 2016.

[22] W.-L. Zheng and B.-L. Lu, "A multimodal approach to estimating vigilance using EEG and forehead EOG," *Journal of Neural Engineering*, vol. 14, no. 2, p. 026017, 2017.

[23] A. Giusti, C. Zocchi, and A. Rovetta, "A noninvasive system for evaluating driver vigilance level examining both physiological and mechanical data," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 1, pp. 127–134, 2009.

[24] M. Patel, S. K. Lal, D. Kavanagh, and P. Rossiter, "Applying neural network analysis on heart rate variability data to assess driver fatigue," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7235–7242, 2011.

[25] J. Meng, B. Zhao, Y. Ma, Y. Ji, and B. Nie, "Effects of fatigue on the physiological parameters of labor employees," *Natural Hazards*, vol. 74, no. 2, pp. 1127–1140, 2014.

[26] M. B. Kurt, N. Sezgin, M. Akin, G. Kirbas, and M. Bayram, "The ANN-based computing of drowsy level," *Expert Systems with Applications*, vol. 36, no. 2, pp. 2534–2542, 2009.

[27] L. Boon-Leng, L. Dae-Seok, and L. Boon-Giin, "Mobile-based wearable-type of driver fatigue detection by GSR and EMG," in *IEEE Region 10 Conf. (TENCON)*, pp. 1–4, 2015.

[28] N. Alioua, A. Amine, and M. Rziza, "Driver's fatigue detection based on yawning extraction," *Int. J. of Veh. Tech.*, vol. 2014, no. 1, pp. 47–75, 2014.

[29] X. Fan, B.-C. Yin, and Y.-F. Sun, "Yawning detection for monitoring driver fatigue," in *Int. Conf. on Machine Learning and Cybernetnics (ICMLC)*, vol. 2, pp. 664–668, 2007.

[30] D. P. Tripathi and N. Rath, "A novel approach to solve drowsy driver problem by using eye-localization technique using cht," *International Journal of Recent Trends in Engineering*, vol. 2, no. 2, p. 139, 2009.

[31] M. Flores, J. Armingol, and A. de la Escalera, "Driver drowsiness warning system using visual information for both diurnal and nocturnal illumination conditions," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. 1, p. 438205, 2010.

[32] T. DąŕOrazio, M. Leo, C. Guaragnella, and A. Distante, "A visual approach for driver inattention detection," *Pattern Recognition*, vol. 40, no. 8, pp. 2341–2355, 2007.

[33] N. Li and C. Busso, "Analysis of facial features of drivers under cognitive and visual distractions," in *IEEE Int. Conf. on Multimedia and Expo (ICME)*, pp. 1–6, 2013.

[34] E. Murphy-Chutorian and M. M. Trivedi, "Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 2, pp. 300–311, 2010.

[35] E. Hoddes, V. Zarcone, H. Smythe, R. Phillips, and W. Dement, "Quantification of sleepiness: a new approach," *Psychophysiology*, vol. 10, no. 4, pp. 431–436, 1973.

[36] T. H. Monk, "A visual analogue scale technique to measure global vigor and affect," *Psychiatry Research*, vol. 27, no. 1, pp. 89–99, 1989.

[37] M. W. Johns, "A new method for measuring daytime sleepiness: the Epworth sleepiness scale," *Sleep*, vol. 14, no. 6, pp. 540–545, 1991.

[38] K. Kaida, M. Takahashi, T. Åkerstedt, A. Nakata, Y. Otsuka, T. Haratani, and K. Fukasawa, "Validation of the karolinska sleepiness scale against performance and EEG variables," *Clinical Neurophysiology*, vol. 117, no. 7, pp. 1574–1581, 2006.

[39] C. C. Liu, S. G. Hosking, and M. G. Lenné, "Predicting driver drowsiness using vehicle measures: Recent insights and future challenges," *Journal of Safety Research*, vol. 40, no. 4, pp. 239–245, 2009.

[40] Q. He, W. Li, and X. Fan, "Estimation of driver's fatigue based on steering wheel angle," in *Int. Conf. on Engineering Psychology and Cognitive Ergonomics (EPEC), Springer*, pp. 145–155, 2011.

[41] D. Das, S. Zhou, and J. D. Lee, "Differentiating alcohol-induced driving behavior using steering wheel signals," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1355–1368, 2012.

[42] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.

[43] Z. Zhou, Y. Wang, Q. J. Wu, C.-N. Yang, and X. Sun, "Effective and efficient global context verification for image copy detection," *IEEE Trans. on Information Forensics and Security*, vol. 12, no. 1, pp. 48–63, 2017.

[44] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[45] B. Gu, V. S. Sheng, K. Y. Tay, W. Romano, and S. Li, "Incremental support vector learning for ordinal regression," *IEEE Trans. on Neur. net. and lear. syst.*, vol. 26, no. 7, pp. 1403–1416, 2015.

[46] T. Akilan, Q. M. J. Wu, A. Safaei, and W. Jiang, "A late fusion approach for harnessing multi-cnn model high-level features," in *2017 IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC)*, pp. 566–571, 2017.

[47] T. Akilan, Q. M. J. Wu, Y. Yang, and A. Safaei, "Fusion of transfer learning features and its application in image classification," in *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1–5, 2017.

[48] Y. Yang and Q. J. Wu, "Extreme learning machine with subnetwork hidden nodes for regression and classification," *IEEE Trans. Cybernetics*, vol. 46, no. 12, pp. 2885–2898, 2016.

[49] B. Schuller, "Recognizing affect from linguistic information in 3d continuous space," *IEEE Trans. Affect. Comput*, vol. 2, no. 4, pp. 192–205, 2011.

[50] Y. Dong, S. Gao, K. Tao, J. Liu, and H. Wang, "Performance evaluation of early and late fusion methods for generic semantics indexing," *Pattern Analysis and Applications*, vol. 17, no. 1, pp. 37–50, 2014.

[51] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.

[52] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, *et al.*, "Going deeper with convolutions," Cvpr, 2015.

[53] A. Delorme and S. Makeig, "Eeglab: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.

[54] C.-T. Lin, L.-W. Ko, I.-F. Chung, T.-Y. Huang, Y.-C. Chen, T.-P. Jung, and S.-F. Liang, "Adaptive EEG-based alertness estimation system by using ICA-based fuzzy neural networks," *IEEE Trans. Circuits-I*, vol. 53, no. 11, pp. 2469–2476, 2006.

[55] Z. Koldovsky, P. Tichavsky, and E. Oja, "Efficient variant of algorithm fastica for independent component analysis attaining the cramér-rao lower bound," *IEEE Trans. on neural networks*, vol. 17, no. 5, pp. 1265–1277, 2006.

[56] L.-C. Shi, Y.-Y. Jiao, and B.-L. Lu, "Differential entropy feature for EEG-based vigilance estimation," in *Engineering in Medicine and Biology Society (EMBC), 35th Int. IEEE Annu. Conf.*, pp. 6627–6630, 2013.

[57] D. F. Dinges and R. Grace, "Perclos: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance," *US Department of Transportation, Federal Highway Administration, Publication Number FHWA-MCRT-98-006*, 1998.

[58] J.-F. Xie, M. Xie, and W. Zhu, "Driver fatigue detection based on head gesture and perclos," in *IEEE Int. Conf. on Wavelet Active Media Technology and Information Processing (ICWAMTIP)*, pp. 128–131, 2012.

[59] D. Sommer and M. Golz, "Evaluation of perclos based current fatigue monitoring technologies," in *IEEE Int. Conf. on Engineering in Medicine and Biology Society (EMBC)*, pp. 4456–4459, 2010.

[60] M. Ferrara and L. De Gennaro, "How much sleep do we need?," *Sleep Medicine Reviews*, vol. 5, no. 2, pp. 155–179, 2001.

[61] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 2, pp. 513–529, 2012.

[62] Y. Yang, Y. Wang, and X. Yuan, "Bidirectional extreme learning machine for regression problem and its learning effectiveness," *IEEE Trans. Neur. Net. Lear. Syst.*, vol. 23, no. 9, pp. 1498–1505, 2012.

[63] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.

[64] T. Baltrušaitis, N. Banda, and P. Robinson, "Dimensional affect recognition using continuous conditional random fields," in *Automatic Face and Gesture Recognition (FG), 10th IEEE Int. Conf. Workshops*, pp. 1–8, 2013.

[65] V. Imbrasaitė, T. Baltrušaitis, and P. Robinson, "Ccnf for continuous emotion tracking in music: Comparison with ccrf and relative feature representation," in *Multimedia and Expo Workshops (ICMEW), IEEE Int. Conf.*, pp. 1–6, 2014.

[66] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," 1999.

[67] Y. Yang, Q. J. Wu, and Y. Wang, "Autoencoder with invertible functions for dimension reduction and image reconstruction," *IEEE Trans. on Syst., Man, and Cy. (TSMC): S.*, vol. 48, no. 7, pp. 1065–1079, 2018.

Wei Wu (S'17) received his M.S. degree in electrical engineering from Hunan University of Technology, China, in 2011.

He was with the College of Electrical and Information Engineering, Hunan University of technology (04-08, and 11-16), China. He is currently pursuing the Ph.D. degree in electrical engineering with the College of Electrical and Information Engineering, Hunan University, Changsha, China. He has been awarded a scholarship under the State Scholarship Fund by the China Scholarship Council (CSC) to continue pursuing his Ph.D. degree with the Department of Electrical and Computer Engineering, University of Windsor, ON, Canada. His research focuses on machine learning, neural networks, and affective computing.

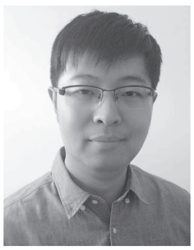
Q. M. Jonathan Wu (M'92-SM'09) received his Ph.D. in electrical engineering from the University of Wales, Swansea, U.K., in 1990.

He was affiliated with the National Research Council of Canada for ten years beginning in 1995, where he became a Senior Research Officer and a Group Leader. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada. He is a Visiting Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. He has published more than 300 peer-reviewed papers in computer vision, image processing, intelligent systems, robotics, and integrated microsystems. His current research interests include 3-D computer vision, active video object tracking and extraction, interactive multimedia, sensor analysis and fusion, and visual sensor networks.

Dr. Wu holds the Tier 1 Canada Research Chair in Automotive Sensors and Information Systems. He is an Associate Editor of the IEEE Transactions on Neural Networks and Learning Systems, and the Cognitive Computation. He has served on technical program committees and international advisory committees for many prestigious conferences.

Wei Sun received his B.S, M.S, and Ph.D. degrees from the Department of Control Science and Engineering, Hunan University, P. R. China, in 1996, 1999 and 2003, respectively.

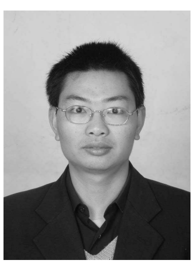
He is now working as a Professor at the College of Electrical and Information Engineering, Hunan University. He also is the director of Hunan Provincial Key Laboratory of Intelligent Robot Technology in Electronic Manufacturing. His areas of interests are computer vision and robotics, neural networks, and intelligent control.



Yimin Yang (S'10-M'13) received his Ph.D. degrees in electrical engineering from the College of Electrical and Information Engineering, Hunan University, Changsha, China, in 2013.

From 2014 to 2018, he was a Post-Doctoral Fellow at the University of Windsor, ON., Canada. He is currently an Assistant Professor with the Department of Computer Science at the Lakehead University, ON., Canada. His research interests are artificial neural networks, signal processing, and robotics.

Dr. Yang was the recipient of the Outstanding Ph.D. Thesis Award of Hunan Province, and the Outstanding Ph.D. Thesis Award Nominations of Chinese Association of Automation, China, in 2014 and 2015, respectively. He has been serving as a reviewer for international journals of his research field, a Guest Editor of multiple journals, and a Program Committee Member of some international conferences.



Xiaofang Yuan (M'14) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Hunan University, Changsha, China, in 2001, 2006, and 2008, respectively.

In 2008, he joined the College of Electrical and Information Engineering, Hunan University. He is currently a Professor with the College of Electrical and Information Engineering, Hunan University. His research interests include artificial neural networks, industrial process control, and computing intelligence.

Wei-Long Zheng (S'14) received the bachelor's degree in information engineering from the Department of Electronic and Information Engineering, South China University of Technology, Guangzhou, China, in 2012, and the Ph.D. degree in computer science from the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China, in 2018.

His research focuses on affective computing, brain-computer interface, machine learning, and pattern recognition. He obtained the IEEE Transactions on Autonomous Mental Development Outstanding Paper Award in 2018.

Bao-Liang Lu (M'94-SM'01) received the B.S. degree in instrument and control engineering from Qingdao University of Science and Technology, Qingdao, China, in 1982, the M.S. degree in computer science and technology from Northwestern Polytechnical University, Xian, China, in 1989, and the Dr. Eng. degree in electrical engineering from Kyoto University, Kyoto, Japan, in 1994.

He was with Qingdao University of Science and Technology from 1982 to 1986. From 1994 to 1999, he was a Frontier Researcher with the Bio-Mimetic Control Research Center, Institute of Physical and Chemical Research (RIKEN), Nagoya, Japan, and a Research Scientist with the RIKEN Brain Science Institute, Wako, Japan, from 1999 to 2002. Since 2002, he has been a Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. He has also been an Adjunct Professor with the Laboratory for Computational Biology, Shanghai Center for Systems Biomedicine, since 2005. His current research interests include brain-like computing, neural network, machine learning, computer vision, bioinformatics, brain-computer interface, and affective computing.

Dr. Lu was the President of the Asia Pacific Neural Network Assembly (APNNA) and the General Chair of the 18th International Conference on Neural Information Processing in 2011. He is currently an Associate Editor of the IEEE Transactions on Cognitive and Developmental Systems, the Neural Networks and a Board Member of APNNA.